

Problem Description

Visitors are attracted to a website through a variety of marketing channels, including the following:

- **video:** Video advertisements; e.g., YouTube
- **organic search:** Unpaid search engine results
- **paid search:** Paid advertisements on search engines
- **display:** Display advertisements embedded within web pages
- **partner referral:** Referrals from other applications that integrate with our products; e.g., Google Docs
- **social:** Advertisements run in social media; e.g., LinkedIn
- **email:** Email advertisements; e.g., in Gmail

These channels redirect the user to a "landing page," a web page designed to showcase value and convince visitors to start a subscription or trial, at which point they receive access to the product.

The marketing team has been tracking site visits and subsequent subscriptions for two years now, but, up until now, no one has had a chance to thoroughly explore and understand the data. That's where you come in! We suspect that certain customer segments are more valuable than others, but we need your help to understand those differences and develop recommendations for improvement.

Data Description

You'll be working with data from two tables: `visits.csv` and `subscriptions.csv`.

The "visits" data set records each account's first visit to our website. This means that the data set contains only one visit per account. For the purpose of this assignment, you can treat "visits" and "visitors" as the same thing. This data source includes the following fields:

- **visit_id:** Unique identifier for the visit
- **account_id:** Unique identifier for the account created by this visit;
note: an account is only created when a visitor starts a subscription or a trial
- **day:** Date on which the visit occurred
- **region:** Geographic region associated with the visit
- **language:** Browser language associated with the visit
- **channel:** Marketing channel responsible for generating the visit
- **landing_page:** Landing page visited;
potential values include the capital letters A through J

The "subscriptions" data set contains an entry for the first subscription (if any) generated from each site visit. It contains the following fields:

- **subscription_id:** Unique identifier for this subscription
- **account_id:** Same as the account_id in the visits table, except this table should contain an account_id in every row
- **subscription_start_date:** Start date for the subscription, if applicable
- **subscription_end_date:** End date for the subscription, if applicable
- **trial_start_date:** Start date for the trial, if applicable
- **trial_end_date:** End date for the trial, if applicable
- **revenue:** Revenue resulting from the subscription in USD;
note: Trials do not generate revenue, so this row is populated for subscriptions only.

Questions

1. How much revenue did each region generate within the time frame covered by the data sets?
2. Which channel had the highest revenue per visit? What was its revenue per visit?
3. Which landing page had the highest overall conversion rate (% of visitors that go on to become paying subscribers)? What was its conversion rate?
4. Did any regions have a different top converting landing page (as compared to the previous answer)? If so, please list them along with their top landing page and its conversion rate.
5. Which region had the highest direct subscription rate (% of visitors that start a subscription directly without doing a trial)? What was its direct subscription rate?
6. Which region had the highest trial start rate (% of visitors that start a trial, regardless of whether those users ultimately cancel or become paying subscribers)? What was its trial start rate?
7. Which region had the highest trial conversion rate (% of trials that start a subscription)? What was its trial conversion rate?
8. Suppose that the company has a goal to increase revenue from the US/Canada region by 20% over the next year. Provide at least 4 potential strategies we could employ to help us reach our goal (no supporting data analysis required - just ideas).